

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2002, 07:03:59 ; Search time 28.35 Seconds  
(without alignments)  
699.671 Million cell updates/sec

Title: US-09-652-292-2

Perfect score: 2765

Sequence: 1 MGHSPVLPVLCASVSLGL.....GHRNCTGIPYSRIEISAAS 541

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	592	21.4	461	1 CSBC_BACSU	P46333 bacillus su
2	571	20.7	457	1 XYL2_LACBR	O52733 lactobacilli
3	520.5	18.8	557	1 ITR2_SCHPO	P87110 schizosacch
4	499	18.0	464	1 GALP_ECOLI	P37021 escherichia
5	497.5	18.0	547	1 GTR1_LEIDO	Q01440 leishmania
6	495	17.9	472	1 ARAE_ECOLI	P09830 escherichia
7	489.5	17.7	472	1 ARAE_KLEOX	P45598 klebsiella
8	484.5	17.5	522	1 STPI_ARATH	P23586 arabidopsis
9	473.5	17.1	575	1 ITR1_SCHPO	Q10286 schizosacch
10	467	16.9	523	1 STC_RICCO	Q41144 ricinus com
11	464.5	16.8	468	1 GLCP_SYNY3	P15729 synecocyst
12	460	16.6	491	1 XYLE_ECOLI	P09098 escherichia
13	452.5	16.4	482	1 YFIC_BACSU	P54723 bacillus su
14	438	15.8	584	1 ITR1_YEAST	P30605 saccharomyc
15	430.5	15.6	522	1 STPI_RICCO	Q10710 ricinus com
16	418.5	15.1	522	1 ITR2_RAT	P12336 rattus norv
17	409	14.8	612	1 ITR2_YEAST	P30606 saccharomyc
18	408.5	14.8	473	1 GLP_ZYMMO	P21906 zymomonas m
19	407.5	14.7	509	1 GTR4_HUMAN	P14672 homo sapien
20	405.5	14.7	495	1 GTR3_CANFA	P47842 canis famil
21	405	14.6	510	1 GTR4_MOUSE	P14142 mus musculu
22	404.5	14.6	494	1 GTR3_SHEEP	P47843 ovis aries
23	403	14.6	523	1 GTR2_MOUSE	P14246 mus musculu
24	401.5	14.5	509	1 GTR4_RAT	P19357 rattus norv
25	399	14.4	763	1 RGT2_YEAST	Q12300 saccharomyc
26	397	14.4	546	1 HXT0_YEAST	P43581 saccharomyc
27	396	14.3	510	1 HXPC_RICCO	Q07423 ricinus com
28	396	14.3	567	1 RAG1_KLULA	P18631 kluyveromyc
29	395.5	14.3	540	1 HUP2_CHLKE	Q39524 chlorella k
30	394.5	14.3	496	1 GTR3_CHICK	P28568 gallus gall
31	394.5	14.3	524	1 GTR2_HUMAN	P11168 homo sapien
32	392.5	14.2	534	1 HUP1_CHLKE	P15686 chlorella k
33	390	14.1	541	1 HXT2_YEAST	P23585 saccharomyc

34	388	14.0	533	1 GTR2_CHICK	Q90592 gallus gall
35	387	14.0	534	1 HUP3_CHLKE	Q39525 chlorella k
36	383.5	13.9	490	1 GTR1_CHICK	P46896 gallus gall
37	381	13.8	551	1 HGT1_KLULA	P49374 kluyveromyc
38	381	13.8	594	1 RCO3_NEUCR	Q92253 neurospora
39	379.5	13.7	496	1 GTR3_HUMAN	P11169 homo sapien
40	378.5	13.7	493	1 GTR3_MOUSE	P20237 mus musculu
41	376.5	13.6	493	1 GTR3_RAT	Q07647 rattus norv
42	375.5	13.6	492	1 GTR1_HUMAN	P11166 homo sapien
43	375.5	13.6	566	1 HXT2_KLULA	P53387 kluyveromyc
44	374.5	13.5	492	1 GTR1_BOVIN	P27674 bos taurus
45	374.5	13.5	592	1 HXT5_YEAST	P38695 saccharomyc

#### ALIGNMENTS

##### RESULT 1

ID	CSBC_BACSU	STANDARD;	PRT;	461 AA.
AC	P46333; Q32289;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	PROBABLE METABOLITE TRANSPORT PROTEIN CSBC.			
GN	CSBC OR SS92BR.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168 / BGSC1A1;			
RX	MEDLINE=9603926; PubMed=7584049;			
RA	Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;			
RT	*Cloning and sequencing of a 36-kb region of the Bacillus subtilis			
RT	genome between the gnt and lol operons.*;			
RL	DNA Res. 2:61-69(1995).			
[2]				
RP	REVISIONS.			
RA	Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;			
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.			
[3]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RA	Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;			
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.			
[4]				
RN	CHARACTERIZATION.			
RX	MEDLINE=99303315; PubMed=10376822;			
RA	Akbar S., Lee S.Y., Roylan S.A., Price C.W.;			
RT	*Two genes from Bacillus subtilis under the sole control of the			
RT	general stress transcription factor sigmaab.*;			
RL	Microbiology 145:1069-1078(1999).			
CC	-1- FUNCTION: COULD SERVE EITHER A NUTRITIONAL OR AN OSMOTIC			
CC	PROTECTION FUNCTION.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).			
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB005554; BAA21604.1; -			
DR	EMBL; Z99124; CAB16017.1; -			
DR	Subtilist; BG11360; csbc.			
DR	InterPro; IPR003663; Sugar_transportr.			
DR	InterPro; IPR003662; sub_transportr.			
DR	Pfam; PF00083; sugar_tr; 1.			

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
OC Lactobacillus.  
OX NCBI\_TaxID=1580;



RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.,  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RL Science 277:1453-1474(1997).  
 CC -1- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH  
 CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.  
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U28377; AAA69110.1; -;  
 DR EMBL; AE000377; AAC75980.1; -;  
 DR EcoGene; EG12148; galP.  
 DR InterPro; IPR003663; Sugar\_transportr.  
 DR InterPro; IPR003662; sub\_transportr.  
 DR Pfam; PF00083; sugar\_tr; 1.  
 DR PRINTS; PR00171; SUGTRNSPORT.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 DR Transport; Sugar transport; Transmembrane; Inner membrane; Symport;  
 KW Complete proteome.  
 FT TRANSMEM 16 36 POTENTIAL.  
 FT TRANSMEM 57 77 POTENTIAL.  
 FT TRANSMEM 85 105 POTENTIAL.  
 FT TRANSMEM 113 133 POTENTIAL.  
 FT TRANSMEM 140 160 POTENTIAL.  
 FT TRANSMEM 172 192 POTENTIAL.  
 FT TRANSMEM 251 271 POTENTIAL.  
 FT TRANSMEM 291 311 POTENTIAL.  
 FT TRANSMEM 322 342 POTENTIAL.  
 FT TRANSMEM 352 372 POTENTIAL.  
 FT TRANSMEM 395 415 POTENTIAL.  
 FT TRANSMEM 417 437 POTENTIAL.  
 SQ SEQUENCE 464 AA; 50982 MW; 07E08935BD8E3F8E CRC64;

Query Match 18.0%; Score 499; DB 1; Length 464;  
 Best Local Similarity 27.2%; Pred. No. 3.2e-24;  
 Matches 141; Conservative 95; Mismatches 183; Indels 100; Gaps 10;

QY 14 VSLGGTFCGYELAVISGALLPLQIDFGLSCLEQDFLYGSLGALLGALLASLVGLFIDCYG 73  
 Db 21 LAALAGLLFGLDGVAGALPFIADFOITSTQEWVSSMMFGAAGVAGSGWLSFKLG 80

QY 74 RKQAILGSNLVLLAGSLTLGLAGSLAVLGRVVGFAISLSMMACCIYVSELVGPORC 133  
 Db 81 RKSLMIGALFLVAGSLFSAAPNVEVILSRVLLGLAVGVASYTAPDLYSEIAPEKIRG 140

QY 134 VLVSLEYAGITVGLISYALNYALAGTPWGRHFMFGWATAPAVLQSLSLFLPAGTDETA 193  
 Db 141 SMISMVQLMITIGILGAYLSDTAPSYT-GANRWMLGVIIIPAILLIGVFFLPDPSRWFA 199

QY 194 THKDLPLQ-----GGEAPK-LGPGRP-----RYSFLDLFRARDNMGRRTVGLGLV 239  
 Db 200 AKRRFYDAERVLRLRDTSEAKDELDEIRESLQVKGSGWALFKENSFRRAVFLGVLLQ 259

QY 240 LFQQLTGQPNVLCYASTIFSSGVFGHGSSAVLASVGLGAVKVAATLTAMGLVDRGRRL 299  
 Db 260 VHQQFTGMVIVYAKIFELAGYTNTQEWGTVVGLTNVLFATFIAIGLVDRWGRKPT 319

QY 300 LLAGCALMALSVSGIGLVFAPVMDSPGSLAVPNATGQTGLPGDGLLQDSLSLPPIPT 359  
 Db 320 LTLGLFLVMA---AGMGVLG----- 335

QY 360 NEDOREPILSTAKTKPHRPGDPSAPPRALSSALPGPPLPARGHALLRWALLCLMVF 419  
 Db 336 -----TMMHIGIHSPSA-----QYFAIAMLMF 358

QY 420 VSAFSGFGPVTWLVLSEIYPVEIRGAFAP-CN-SFNWAANLFISLSFLDLITIGLSW 477  
 Db 359 IVGFAMAGPLIWLICSELQ--LKRDFGTCSTATNWIANNIVGATFUTMLTNGAN 416

QY 478 TFLYGLTAVLGLGFIYLFVPETKQSLAEIDQOFKRR 516  
 Db 417 TFWYAAALNVLLITLWLVPETKHSVLEHIERNLMKGR 455

RESULT 5  
 GTRL\_LEIDO  
 ID GTRL\_LEIDO STANDARD; PRT; 547 AA.  
 AC Q01440;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DE 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE MEMBRANE TRANSPORTER D1.  
 OS Leishmania donovani.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_Taxid=5661;  
 RP [1]  
 RM SEQUENCE FROM N.A.  
 RX MEDLINE=93063058; PubMed=1435876;  
 RA Langford C.K., Ewhank S.A., Hanson S.S., Ullman B., Landfear S.M.;  
 RT "Molecular characterization of two genes encoding members of the  
 RT glucose transporter superfamily in the parasitic protozoan *Leishmania*  
 RT *donovani*.";  
 RL Mol. Biochem. Parasitol. 55:51-64(1992).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M85072; AAA29230.1; -;  
 DR PIR; A48442; A48442.  
 DR InterPro; IPR003663; Sugar\_transportr.  
 DR InterPro; IPR003662; sub\_transportr.  
 DR Pfam; PF00083; sugar\_tr; 1.  
 DR PRINTS; PR00171; SUGTRNSPORT.  
 DR PRINTS; PR00172; GLUCTRNSPORT.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 KW Transport; Sugar transport; Transmembrane.  
 FT TRANSMEM 3 25 POTENTIAL.  
 FT TRANSMEM 44 64 POTENTIAL.  
 FT TRANSMEM 79 99 POTENTIAL.  
 FT TRANSMEM 102 122 POTENTIAL.  
 FT TRANSMEM 137 157 POTENTIAL.  
 FT TRANSMEM 165 185 POTENTIAL.  
 FT TRANSMEM 246 266 POTENTIAL.  
 FT TRANSMEM 277 297 POTENTIAL.  
 FT TRANSMEM 309 329 POTENTIAL.  
 FT TRANSMEM 340 360 POTENTIAL.  
 FT TRANSMEM 386 406 POTENTIAL.  
 FT TRANSMEM 408 428 POTENTIAL.  
 SQ SEQUENCE 547 AA; 58787 MW; CA8EEF16EB0FF514 CRC64;

Query Match 18.0%; Score 497.5; DB 1; Length 547;  
 Best Local Similarity 26.4%; Pred. No. 4.7e-24;  
 Matches 146; Conservative 83; Mismatches 190; Indels 135; Gaps 15;

QY 10 LCASVSLGGLTFGYELAVISGALLPLQIDFGLS--CLEQEFVLGSLGALLASLVGSGF 67

```
Db 7 LCNA---LGGFLGYDTGVINAALFOKDFGSEHSQYALIVAIAIAGAFVGFAPISGF 63
Qy 68 LIDCYGRKQAILGSLNVLGAGSLWGLVGRVAVGFAISLSSMACCIYSELV 127
Db 64 ISAAEGRRCIAVADALFVIGSLMGAPNVVLSRVIVGLAIGISSATIPVYLAET 123
Qy 128 GPRQGVLSVLEAGITVIGILLS---YALNYALAGTPMGWRHMGWATAPAVLQSLSLF 184
Db 124 SPKRGATVNLNLFITGCGFVAAGFTAIMVFTSKNIGRWVAIGIGALPVAVQACLLF 193
Qy 185 -----LPACTDETAT-----HKDLIPLQ-GGEAPKLGPRPRYSFLDLFRARDNM 228
Db 184 FLPEPSRWLLSKHADRAKAVADKFEVDLCFQEGDELPVS-----RIDYRPLM-ARD-M 236
Qy 229 RGRVTTVGLGLVFOQLTQCPNVLCYASTIFSSVGFHGGSSAVLASVGLCAVKVAATLTAM 288
Db 237 RFVVLSSGLQIIQFSGINTIIMYSSVILYDAGFRDAIMPVVLISPLAFMNAFLTAVAI 296
Qy 289 GLVDAGRRALLIA---GCALMALSVSGIGL-----VSFAVPMDSGSPCLAVPNATGOTG 340
Db 297 FTVDRGRRLMLUISVFGCVLLVIVIAIGFFIGTRISYSV----- 337
Qy 341 LFGDSGLLDQSLPPPIPTNEDQREPIILSTAKTKPRPSGDPSPAPPRALSSALPGPPL 400
Db 338 ---GGGLF----- 342
Qy 401 PARGHALLRWLTALLCLMVVVSFAFGFPGVPMVLVSEIYVEIRGAFACFSFNMAANL 460
Db 343 -----LALLAVFLALYAPGICCPVINGEIEFTHLRTSAASVATMANNGANV 390
Qy 461 FLSFLDLIGTIGTSWTFLLYGLTVGLGFIYLFVDPETKGQSLAEIDQFOKR----- 515
Db 391 LVSVFPILMGALCVGCTFTIISGLMALGCIFFVFAVETKGLTLEQIDNMPKRAGLPP 450
Qy 516 RETLSFGHRONSG 529
Db 451 RF-----HEEGESG 459

RESULT 6
ID ARAE_ECOLI STANDARD; PRT: 472 AA.
AC P09830; Q46937;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ARABINOSE-PROTON SYMPORTER (ARABINOSE TRANSPORTER).
GN ARAE OR B2841 OR 24161 OR ECS3698.
OS Escherichia coli, and
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxId=562, 83334;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87115869; PubMed=3543693;
RA Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
Henderson P.J.F.;
RT "Mammalian and bacterial sugar transport proteins are homologous.";
RL Nature 325:641-643(1987).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / JM2433;
RX MEDLINE-88228015; PubMed=2836407;
RA Maiden M.C.J., Jones-Mortimer M.C., Henderson P.J.F.;
RT "The cloning, DNA sequence, and overexpression of the gene araE
coding for arabinose-proton symport in Escherichia coli K12.";
RL J. Biol. Chem. 263:8003-8010(1988).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed=9278503;
```

```
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE-21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grothbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE-21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takani H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[6]
RP PRELIMINARY SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE-84114868; PubMed=6319708;
RA Stoner C., Schleif R.F.;
RT "The araE low affinity L-arabinose transport promoter. Cloning,
sequence, transcription start site and DNA binding sites of
regulatory proteins.";
RL J. Mol. Biol. 171:369-381(1983).
CC -1- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/
or send an email to license@isb-sib.ch).
CC
CC EMBL; J03732; AAA23469.1; -
CC EMBL; X00272; CAA25075.1; -
CC EMBL; U29581; AAB40488.1; -
CC EMBL; AE000368; AAC75880.1; -
CC EMBL; AE005513; AAG57953.1; -
CC EMBL; AP002563; BAB37121.1; -
CC PIR; B26430; B26430.
CC PIR; A28075; A28075.
CC EcoGene; EGI0056; araE.
CC InterPro; IPR003663; Sugar_trnsportr.
CC InterPro; IPR003662; sub_trnsportr.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PRO0171; SUGTRNSPORT.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
Complete proteome.
FT TRANSMEM 23 43 POTENTIAL.
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 92 110 POTENTIAL.
FT TRANSMEM 120 138 POTENTIAL.
FT TRANSMEM 149 167 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
```

```

FT TRANSMEM 259 278 POTENTIAL.
FT TRANSMEM 299 318 POTENTIAL.
FT TRANSMEM 328 348 POTENTIAL.
FT TRANSMEM 362 382 POTENTIAL.
FT TRANSMEM 396 415 POTENTIAL.
FT TRANSMEM 424 443 POTENTIAL.
FT CONFLICT 26 28 SVA -> YDR (IN REF. 6).
SQ SEQUENCE 472 AA; 51684 MW; 411990A441D44393 CRC64;

Query Match 17.9%; Score 495; DB 1; Length 472;
Best Local Similarity 25.3%; Pred. No. 5.7e-24;
Matches 137; Conservative 93; Mismatches 182; Indels 132; Gaps 8;

QY 17 LGGLTFGEYELAVISGALLPLQDPLGSLCEQEFVLGSLLLGALLASVGGFLIDCYGRKQ 76
Db 31 VAGLFGDIDGIVAGALPITDHFVLTSLQEWVSSMMLGAALGALFNGWLSFRLGRKY 90

QY 77 AILGSLNVLGSLTGLAGSLAWLVGRAVGVFAISLSSMACCIYVSELVGPQRGVLV 136
Db 91 SIMAGAILFVLGSGSATSVEMLIAARVVLGIVAGIASYAPLYLSEMASENVVRGKMI 150

QY 137 SLYEAGITVGLLSYALNYALAGTPWGRHMFGWATAPAVLQSLSLFLP----- 186
Db 151 SMYQLMVTGLVIAFLSDTAFSYS-GNWRAMLGVLAIPAVVLLILVFLPNSPRWLAEGK 209

QY 187 -----AGTDETA-----THKDLIPLQGEAPKLGPRPRYSFLDLFRANDM 228
Db 210 RHIEAEVIRMLRDTSEKAREELNEIRESLKQCGWA-----LFKINRV 255

QY 229 RGRITVGLGLVFOOLTGPVNLVYASTIFSSVGFHGGSSAVLASVGLGAVKVAATLTAM 288
Db 256 RRAVFLGMLLQAMQOFTGNNIIMYAPRIKFMAGFTTTEQOMIATVVLGTFEATFIIV 315

QY 289 GLVDRAGRALLAGCALMALS--VSGIGLVSAFVPMDSGSPCLAVPNATGQTLPGDSG 346
Db 316 FTVDKAGRPALKIGFSVMAIGTLVGLCYLMOF-----DNGTASSG----- 356

QY 347 LLDQSSLPPIPTNEDQREPILSTAKTKPHRSGDPSAPPRLLSSALPGPPLPARGHA 406
Db 357 ----- 356

QY 407 LLRTWALLCLMVFSAFSGFGPVTVLVLSEIYPVEIRGAFAPFCNSFWNAANLFLISLF 466
Db 357 ----- 356

QY 467 LDLTGIGLSTWFLYGLTAVLGLGFIYLVFVETKQSLAEIDDOQKRFRTLSFGHQON 526
Db 416 LTLDSIGAAGTFWLYTALNIAFVITWLPETKNVLEHIERK-----LMAGEKLR 468

QY 527 STGI 530
Db 469 NIGV 472

RESULT 7
ARAE_KLEOX
ID ARAE_KLEOX STANDARD; PRT; 472 AA.
AC P45598;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ARABINOSE-PROTON SYMPORTER (ARABINOSE TRANSPORTER).
GN ARAB.
OS Klebsiella oxytoca.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Klebsiella.
NCBI_TaxID=571;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8017;
RX MEDLINE=95394866; PubMed=7665532;
RA Shatwell K.P., Charalambous B.M., McDonald T.P., Henderson P.J.F.;

```

```

RT *Cloning, sequencing, and expression of the araE gene of Klebsiella
RT oxytoca 8017, which encodes arabinose-H+ symport activity.*;
RL J. Bacteriol. 177:5379-5380(1995).
CC -|- FUNCTION: UPTAKE OF ARABINOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X79598; CAA56110.1;
DR InterPro; IPR003663; Sugar_transportr.
DR InterPro; IPR003662; sub_transportr.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport.
FT TRANSMEM 23 43 POTENTIAL.
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 92 110 POTENTIAL.
FT TRANSMEM 120 138 POTENTIAL.
FT TRANSMEM 149 167 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 259 278 POTENTIAL.
FT TRANSMEM 299 318 POTENTIAL.
FT TRANSMEM 328 348 POTENTIAL.
FT TRANSMEM 362 382 POTENTIAL.
FT TRANSMEM 396 415 POTENTIAL.
FT TRANSMEM 424 443 POTENTIAL.
SQ SEQUENCE 472 AA; 51732 MW; 410021E1BEE3D96E CRC64;

Query Match 17.7%; Score 489.5; DB 1; Length 472;
Best Local Similarity 25.3%; Pred. No. 1.3e-23;
Matches 133; Conservative 93; Mismatches 174; Indels 125; Gaps 7;

QY 17 LGGLTFGEYELAVISGALLPLQDPLGSLCEQEFVLGSLLLGALLASVGGFLIDCYGRKQ 76
Db 31 VAGLFGDIDGIVAGALPITDHFVLTSLQEWVSSMMLGAALGALFNGWLSFRLGRKY 90

QY 77 AILGSLNVLGSLTGLAGSLAWLVGRAVGVFAISLSSMACCIYVSELVGPQRGVLV 136
Db 91 SLMVGAFLVAGSVGSATSVEMLLVARIIVLGVAVGIASYTAPLYLSEMASENVVRGKMI 150

QY 137 SLYEAGITVGLLSYALNYALAGTPWGRHMFGWATAPAVLQSLSLFLP----- 186
Db 151 SMYQLMVTGLVIAFLSDTAFSYS-GNWRAMLGVLAIPAVVLLILVFLPNSPRWLAEGK 209

QY 187 -----AGTDETA-----AGTDETA-----AGTDETA----- 228
Db 210 RHVEAEVIRMLRDTSEKAREELNEIRESLKQCGWA-----LFGKINRV 255

QY 229 RGRITVGLGLVFOOLTGPVNLVYASTIFSSVGFHGGSSAVLASVGLGAVKVAATLTAM 288
Db 256 RRAVFLGMLLQAMQOFTGNNIIMYAPRIKFMAGFTTTEQOMIATVVLGTFEATFIIV 315

QY 289 GLVDRAGRALLAGCALMALS--VSGIGLVSAFVPMDSGSPCLAVPNATGQTLPGDSG 346
Db 316 FTVDKAGRPALKIGFSVMAIGTLVGLCYLMOF-----DNGTASSG----- 356

QY 347 LLDQSSLPPIPTNEDQREPILSTAKTKPHRSGDPSAPPRLLSSALPGPPLPARGHA 406
Db 357 ----- 356

QY 407 LLRTWALLCLMVFSAFSGFGPVTVLVLSEIYPVEIRGAFAPFCNSFWNAANLFLISLF 466
Db 357 ----- 356

QY 467 LDLTGIGLSTWFLYGLTAVLGLGFIYLVFVETKQSLAEIDDOQKRFRTLSFGHQON 526
Db 416 LTLDSIGAAGTFWLYTALNIAFVITWLPETKNVLEHIERK-----LMAGEKLR 468

QY 527 STGI 530
Db 469 NIGV 472

```

```

Db 357 -LSWLSVGMTMNCIAGYAMSAAPVWVWILCSEIOPKCRDFGICSTTTTWNWVSNMIICATP 415
QY 467 LDIIGTIGLWTFLLYGLTAVLGLGYFLFVPETKGSIAEIDQO 511
Db 416 LTELDAIGAAGTFLYALNVAFIGVTFWLPETKRYNLTLEHIER 460

RESULT 8
ID STPI_ARATH STANDARD; PRT; 522 AA.
AC P23586;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE GLUCOSE TRANSPORTER (SUGAR CARRIER).
GN STPI.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, LANDSBERG RECTA;
RX MEDLINE=91005995; PubMed=2209537;
RA Sauer N., Friedlaender K., Graeml-Wicke U.;
RT Primary structure, genomic organization and heterologous expression
RT of a glucose transporter from Arabidopsis thaliana.*;
RL EMBO J. 9:3045-3050(1990).
CC -!- FUNCTION: ACTIVE UPTAKE OF HEXOSES. PROBABLE GLUCOSE/HYDROGEN
CC SYMPORT.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; X55350; CAA39037.1; -.
DR PIR; S12042; S12042.
DR InterPro; IPR003663; Sugar_transprtr.
DR InterPro; IPR003662; sub.transprtr.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR PRINTS; PR00172; GLUCTRNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Transport; Sugar transport; Symport.
FT DOMAIN 1 21 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 84 104 POTENTIAL.
FT TRANSMEM 113 133 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 285 305 POTENTIAL.
FT TRANSMEM 321 341 POTENTIAL.
FT TRANSMEM 349 369 POTENTIAL.
FT TRANSMEM 385 405 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT DOMAIN 474 522 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 522 AA; 57596 MW; 68A6C72AFED90380 CRC64;

Query Match 17.5%; Score 484.5; DB 1; Length 522;
Best Local Similarity 26.0%; Pred. No. 2.9e-23;
Matches 154; Conservative 96; Mismatches 183; Indels 159; Gaps 17;

```

```

QY 4 SPVPLPCASVSLGLTFFGYELAVISCAL-LP--LQDFGLSCLQDE----- 48
Db 20 TPEVLETCV-VAAMGGLIFGYDIGISGVTSMPLKREFFPSYVRKQOEDASTNQYCYD 78
QY 49 ----FLVGSILLGALLASLVGGFLIDCYGRKOAILGSLNVLGSLTGLAGSLAWLVL 103
Db 79 SPTLTMTPTSSLYLAALISLVASTVTRKFGRLSMLFGGILFCAGALINGFAKHVWMLIV 138
QY 104 GRAVVGFAISLSSMACCIYVSELVGPGRGVLSVLYEAGITVGLISLYALNVALAGTP-- 161
Db 139 GRILLGGIGCFANQAVPLYLEMAPYKYRGALNIGFOLSITIGLVAEVLNLYFFAKIKGG 198
QY 162 WGRHMFGWATAPAVLOSLSLLFLP-----AGTDEATHAKDLI 199
Db 199 WGNRLSLGGAAVVPALIIITIGSLVLPDTPNSMIERGQHEBAKTKLRRIRGVDDVSQERDDL 258
QY 200 PLOGGAPKLGPRPRYSFIDLFRADNWRGRTTVGLGLVLFOOLTGPVNLVYASTIFS 259
Db 259 VAASKESQSI-----EHPWRNLLRRK--YRPHLTMAVMIPFPQOLTGINVIMFYAPVLFN 311
QY 260 SVGFHGGSSAVLASVGLGAVKVAATLTAMGLVDRAGRALLLAGCALMALSVSGLVSP 319
Db 312 TIGFTTDS-LMSAVVTGSVNVGATLVSYGVDRGRRLFLFEGGTOMLI----- 360
QY 320 AVPMDSGPSCLAVPNAT-----GOTGLPGDGLLDSSLPPIPTNEDOREPTLSTAKKT 374
Db 361 -----COAVVACIGAKFGVDGTFGE----- 381
QY 375 KPHPRGDPSPAPRLALSSALPGPLPARGHALLRWLTALLC---LMVFVSASFSGFGPVT 431
Db 382 -----LPRKYAIVVVTFCIYVAGFAWSWGPLG 409
QY 432 WLVLSEIYPVEIRGAFACNSFNAAANFLISLFLDLGTI--GLSWTFLLYGLTAVLG 489
Db 410 WLVPSEIFPLEIRSAQSITVSVNNIFITIAQIFLTMCLHKLFGI---PLVFAFFVVM 466
QY 490 LGFIYLFVPETKGSIAEIDQOFQR-----RPF-----TISFGHQRNSTG 529
Db 467 SIFVYIFLPETKGIPIEMGQWRSHWYMSFEVDEGEYGNALPMGNKNSNOAG 518

RESULT 9
ID ITRI_SCHPO STANDARD; PRT; 575 AA.
AC Q10286;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYO-INOSITOL TRANSPORTER 1.
GN ITRI OR SPAC7D4.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 H90;
RX MEDLINE=98228265; PubMed=9560432;
RA Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,
RA Rusu M., Politelea M., Edenharter L., Schweingruber M.E.;
RT Exogenous inositol and genes responsible for inositol transport are
RT required for mating and sporulation in Schizosaccharomyces pombe.*;
RL Curr. Genet. 33:255-261(1998).
RN [2]
RP SEQUENCE OF 290-575 FROM N.A.
RC STRAIN=972;
RA Gentiles S., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----

```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X98622; CAA67211.1; -;  
DR EMBL; Z99532; CAB6718.1; -;  
DR InterPro: IPR003663; Sugar\_trnsporttr.  
DR InterPro: IPR003662; sub\_trnsporttr.  
DR Pfam: PF00083; sugar\_tr: 1.  
DR PRINTS; PRO0171; SUGTRNSPORT.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 2.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; FALSE\_NEG.  
KW Transmembrane, sugar transport; Glycoprotein.  
FT TRANSMEM 88 108 POTENTIAL.  
FT TRANSMEM 131 151 POTENTIAL.  
FT TRANSMEM 158 178 POTENTIAL.  
FT TRANSMEM 188 208 POTENTIAL.  
FT TRANSMEM 217 237 POTENTIAL.  
FT TRANSMEM 248 268 POTENTIAL.  
FT TRANSMEM 351 371 POTENTIAL.  
FT TRANSMEM 378 398 POTENTIAL.  
FT TRANSMEM 402 422 POTENTIAL.  
FT TRANSMEM 440 460 POTENTIAL.  
FT TRANSMEM 485 505 POTENTIAL.  
FT TRANSMEM 509 529 POTENTIAL.  
FT DOMAIN 530 575 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 575 AA; 62757 MW; 3B7C5EPF86C596AE CRC64;

Query Match 17.1%; Score 473.5; DB 1; Length 575;  
Best Local Similarity 27.0%; Pred. No. 1.5e-22;  
Matches 151; Conservative 88; Mismatches 187; Indels 133; Gaps 16;

QY 7 VLPLCASVSLGLGTFYELAVISGALLPLQLDFG---LSCLQEQLVGLSLLGALLASL 63  
DB 89 VLFAAGI---GGLFGYDTGVISGALVVGTSGLGHELTNGKFEITSATSLGALLGI 145  
QY 64 VGGFLIDCYGRKQAILGSLNVLGASLTGLAGSLAWLVGRAVVGFAISLSSMACCTIV 123  
DB 146 IAGALADPFRKPVIAIASIIIVGSIVQVTAHLWHMIVGREVIGVGVASLIIPYL 205  
QY 124 SELVGRORGLVSLYEAGITVIGLLSYALNYALACTPWGRHMFGWATAPAVLSLSL 183  
DB 206 SEIAPSKIRGLVLIYVLLITAGQVIATYAGIDTAFEHVHNGRMVGLAMVPAFQLFILI 265  
QY 184 FLP-----AGTDETHKDI.I-----PLOGGKAPKLGPG 212  
DB 266 WLPSPRLVYKERSQEAINTLARIYPTAHPIEITKLYIQEGVRDPFGSGRQKIVK- 324  
QY 213 RPRYSFLDLFRARDNRGRTTVGLGLVLFQQLTGQPNVLCYASTIFSSVFGHGSSAVLA 272  
DB 325 -----TFKELYNPNFRA-LILACSLQAMQQLSGNSLMYFSTIEFVGVFNPTATGLI 379  
QY 273 SVGLGAVKVAATLTAMGLVDRAGRALLAGCALMALSVSGI--GLVSPAVPMDSGSPCL 330  
DB 380 ---IAATNFVFTIVAFGVDFPGRILL-----LTWGMIAALIVCAVAFHFLP--- 426  
QY 331 AVPNATGOTGLPGDGLLQDSSLPPIPTNEDOREPILSTAKTKTPHRSQDPSAPRLA 390  
DB 427 -----KDENGNYTSGQSN----- 440  
QY 391 LSSALPGPPLPARGHALLRW--TALLCLMVFVSFAFGFVPTLVLSLIVPEIRGRAF 448  
DB 441 -----NAIVVLSIMVIVVASYASGLNLPW--QQSELPFMSVRGLGT 480

QY 449 AFCNGFNWAANFISLSELDLICTIGLSWTFLLYGLTAVLG-LSFIYLFVPETKQSLAE 507  
DB 481 GNSTAVNAGNIGASFTLMEITPTGTGTFYGLGCLGFLWGLGALFCY-PDLTDYTFIE 539  
QY 508 IDQOFQKRRTLSFGHQRN 526  
DB 540 IGEELKH-----GFGVRES 553

RESULT 10  
STC\_RICCO

ID STC\_RICCO STANDARD; PRT; 523 AA.  
AC Q41144; Q41147;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE SUGAR CARRIER PROTEIN C.  
GN STC.  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV, CARMENCITA; TISSUE=Cotyledon;  
RA Weig A., Franz J., Sauer N., Komor E.;  
RT "Isolation of a family of cDNA-clones from Ricinus communis L.  
with close homology to the hexose carriers";  
RL J. Plant Physiol. 143:178-183(1994).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).  
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L08196; AAA79761.1; -;  
DR EMBL; L08191; AAA79764.1; -;  
DR InterPro: IPR003663; Sugar\_trnsporttr.  
DR InterPro: IPR003662; sub\_trnsporttr.  
DR Pfam: PF00083; sugar\_tr: 1.  
DR PRINTS; PRO0171; SUGTRNSPORT.  
DR PRINTS; PRO0172; GLUCTRNSPORT.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Transmembrane, Transport; Sugar transport; Symport.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 85 105 POTENTIAL.  
FT TRANSMEM 120 140 POTENTIAL.  
FT TRANSMEM 143 163 POTENTIAL.  
FT TRANSMEM 205 225 POTENTIAL.  
FT TRANSMEM 298 318 POTENTIAL.  
FT TRANSMEM 321 341 POTENTIAL.  
FT TRANSMEM 351 371 POTENTIAL.  
FT TRANSMEM 387 407 POTENTIAL.  
FT TRANSMEM 430 450 POTENTIAL.  
FT TRANSMEM 455 475 POTENTIAL.  
SQ SEQUENCE 523 AA; 57769 MW; 074804EFA3A27F6B CRC64;

Query Match 16.9%; Score 467; DB 1; Length 523;  
Best Local Similarity 26.2%; Pred. No. 3.5e-22;  
Matches 148; Conservative 91; Mismatches 161; Indels 164; Gaps 16;

QY 14 VSLGLGTFYELAVISGALLPLQLDFGLSCLQEQL 49  
DB 31 VAAMGGLIFGYDIG-LSGCVTSND-----SFLKKFFPSYRKKKADESSNOYQVDSOTL 84



QY 50 --LVGSLGALLASLVGGFLIDCYGRKQAILGSLNVLVLAGSLTTLGAGSLAWLVIGRAV 107  
 Db 85 TMTSSLYLAALIASLVASTITRFRKLSMLFGVLCAGAI INGAARAVWMLILGRIL 144  
 QY 108 VGFALSSMACCIYSELVGPGRGVLSLYEAGITVIGILLSYALNYALAGTP--WGWR 165  
 Db 145 LFGIGFANOSVPLYLSEMAPKYRGALNIGFQLSITIGILVANVLYFFAKIKGGWGR 204  
 QY 166 HMGWATAPAVLOSLSLELPACGTDTATHKDLIPLOGEAPKLGPRR-----YSEL 219  
 Db 205 LSLGGAWPALIITVGSVLVLP-----DTPNSMIBRGQHEEARAHILKRVGVEDVDEFT 258  
 QY 220 LDFRARDN-----MRGRTTVGLGLVLFQOLTQCPNVLCYASTISSVGFH 264  
 Db 259 DLVHASEDSKVEHPWRNLLQRYRPHLSMAIAIFFQOLTGINVIMFYAPVLEDTIGF- 317  
 QY 265 GSSAVLASVGLGAVKVAATLTAMGLVDRAGRRALLAGCALMALSVSGIGLVFAVPM 324  
 Db 318 GSDAALMSAVITGLVNVFATMVSIYGVDKWGRRLFLFEGGVQMLI----- 362  
 QY 325 SGPSCLAVPNAT-----QGTGLPGDSGLLDSSLPPIPTNEDQREPIILSTAKTKPHR 379  
 Db 363 ----COAIVAAICAGKFGVDGAPGD----- 383  
 QY 380 SGDPSPAPRLSALPGPLPARGHALLRNTA---LLCLMVFSAFSGFGPVTWLVLS 436  
 Db 384 -----LPQYAVVVLFCIYVSGFAWSGPGPLGLVPS 416  
 QY 437 EIVPEIRGRAFAFCNSFNAAFLF-----ISLSFLDLIGTI--GLSWTFLYGLTAVLGL 490  
 Db 417 EIFPLEIRSA-----QSVNVSVNMFTEFVVAQVFLIMLCHLAFGL---TIFFSFFVLIMS 469  
 QY 491 GFIVLEVPETKGSABIDQOFK 514  
 Db 470 IFVYFELPETHGPIEMGVQWKQ 493

RESULT 11  
 GLCP\_SYNY3  
 ID GLCP\_SYNY3 STANDARD; PRT; 468 AA.  
 AC P15729;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE GLUCOSE TRANSPORT PROTEIN.  
 GN GTR OR GLCP OR SLL0771.  
 OS Synchocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.  
 OX NCBI\_TaxId=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90014182; PubMed=2507869;  
 RA Zhang C.C., Durand M.C., Jeanjean R., Joset F.;  
 RT "Molecular and genetic analysis of the fructose-glucose transport  
 RT system in the cyanobacterium Synchocystis PCC6803";  
 RL Mol. Microbiol. 3:1221-1229(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91346660; PubMed=2129397;  
 RA Schmetterer G.R.;  
 RT "Sequence conservation among the glucose transporter from the  
 RT cyanobacterium Synchocystis sp. PCC 6803 and mammalian glucose  
 RT transporters";  
 RL Plant Mol. Biol. 14:697-706(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96127529; PubMed=8590279;  
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
 RA Sugita M., Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
 RT region from map positions 64 to 92% of the genome";

RL DNA Res. 2:153-166(1995).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X15988; CAA34119.1; -;  
 DR EMBL; X16472; CAA34492.1; -;  
 DR EMBL; D64000; BAA10117.1; -;  
 DR PIR; S06973; S06973;  
 DR PIR; S10014; S10014;  
 DR InterPro; IPR003663; Sugar\_transprtr.  
 DR InterPro; IPR003662; sub\_transprtr.  
 DR Pfam; PF00083; sugar\_tr; 1.  
 DR PRINTS; PR00171; SUGRTRANSPORT.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 KW Transmembrane; Transport; Sugar transport; Complete proteome.  
 FT DOMAIN 1 21 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 22 42 POTENTIAL.  
 FT DOMAIN 43 54 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 55 75 POTENTIAL.  
 FT DOMAIN 76 87 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 88 108 POTENTIAL.  
 FT DOMAIN 109 116 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 117 137 POTENTIAL.  
 FT DOMAIN 138 151 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 152 172 POTENTIAL.  
 FT DOMAIN 173 186 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 187 207 POTENTIAL.  
 FT DOMAIN 208 265 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 266 286 POTENTIAL.  
 FT DOMAIN 287 303 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 304 324 POTENTIAL.  
 FT DOMAIN 325 339 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 340 360 POTENTIAL.  
 FT DOMAIN 361 378 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 379 399 POTENTIAL.  
 FT DOMAIN 400 409 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 410 430 POTENTIAL.  
 FT DOMAIN 431 436 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 437 457 POTENTIAL.  
 FT DOMAIN 458 468 CYTOPLASMIC (POTENTIAL).  
 FT CONFLICT 319 319 L -> I (IN REF. 1).  
 SQ SEQUENCE 468 AA; 49747 MW; D7EC545C4FB38D22 CRC64;

Query Match 16.8%; Score 464.5; DB 1; Length 468;  
 Best Local Similarity 26.3%; Pred. No. 4.4e-22;  
 Matches 141; Conservative 79; Mismatches 192; Indels 125; Gaps 12;  
 QY 10 LCASVSLGGLTFGYELAVISGALLPLQLDFGLSCLEQEFVLSLILGALLASLVGGFLI 69  
 Db 19 LISGVAALGGFLFGFDATAVINGAAVAAKQHEQTDLLTGLSVSLALLGSALGAFGAGPIA 78  
 QY 70 DCYGRKQAILGSLNVLVLAGSLTIGLAGS-----LAWLVGRAVGVFAISLSMACC 120  
 Db 79 DRHGRK-----TWILAALVLETLSSIGSLPTTIWDFIFWRVLG-----GIGVGAASVIAP 129  
 QY 121 IYVSELVGPQRQGVLSLYEAGITVIGILLSYALNYAL-----AGTPW-----GWRHMF 168  
 Db 130 AYIAEVSPAHLRGLGSLQALIVSGIFIALLSNFWIALMAGSAGNPWLFGAAAWRMF 189  
 QY 169 GNATAPAVLQSLSLFLP-----AGTDETAHKDLIPLOGGEAPK-----LGP 211  
 Db 190 WTELIPALLGVCAFLIPESPRYLVAQGGKAA-AILMKVEGDDVPSRTEELQATVSLD 248

OY 212 GRPRYSFLDFRANRNRGRTRTVGLGLVLFQOLQPNVLCVASTIFFSVGPHGGSSAVL 271  
 DB 249 HKRFRS--DLSRRGGLLPVWICMGLSALQOFVGINVIFYSSVLRWSVGTTEERS--LL 305  
 OY 272 ASVGLGAVKVAAYLTANGLVDRGRALLAGCALMALSVSGIGLVFAVPMDSPSCLA 331  
 DB 306 ITVITGFINLTTLVAFAFDKGRKPLLMGSGTITLGLSVVFGATVNGQ---- 361  
 OY 332 VPNATGOTGLPGDSSLQDSSLPIPRTNEDOREPILSTAKTKPHPRSGDPSAPPRAL 391  
 DB 362 -PVLTAAGI----- 370  
 OY 392 SSALPGPLPARGHALLRWTLALCLMVFSAFSGFGPVTWLVSEIYPVEIRGRAFC 451  
 DB 371 -----TALVTANLYVFSFGSWGPIVWVLLGEMFNKTKRAAALSYA 411  
 OY 452 NSEFNAANLFTISLFDLICTIGLSWTFLLYGLTAVLGLGYLFLVFPETKGSGLAEI 508  
 DB 412 AGVQWIANFIISTFTFPLDVLGPGPAYGLYATSAASIFFIWFVFKETKKTLEQM 468

RESULT 12  
 XYLE\_ECOLI  
 ID XYLE\_ECOLI STANDARD; PRT; 491 AA.  
 AC P09098;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE D-XYLOSE-PROTON SYMPORTER (D-XYLOSE TRANSPORTER).  
 GN XYLE OR B4031 OR Z5629 OR EGS5014.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562, 83334;  
 RN {1}  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=88007632; PubMed=2820984;  
 RA Davis E.O., Henderson P.J.F.;  
 RT "the cloning and DNA sequence of the gene xyle for xylose-proton  
 RT symport in Escherichia coli K12.";  
 RL J. Biol. Chem. 262:13928-13932(1987).  
 RN {2}  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=8715869; PubMed=3543693;  
 RA Malden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,  
 RA Henderson P.J.F.;  
 RT "Mammalian and bacterial sugar transport proteins are homologous.";  
 RL Nature 325:641-643(1987).  
 RN {3}  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=94089392; PubMed=8265357;  
 RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,  
 RA Daniels D.L.;  
 RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the  
 RT region from 89.2 to 92.8 minutes.";  
 RL Nucleic Acids Res. 21:5408-5417(1993).  
 RN {4}  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis J.N., Lim A., Dimalanta E.T., Potamoudis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN {5}  
 RN SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 RN {6}  
 RN SEQUENCE OF 1-192 FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=88234001; PubMed=2836810;  
 RA Francoz E., Dassa E.;  
 RT "3' end of the malFG operon in E. coli: localization of the  
 RT transcription termination site.";  
 RL Nucleic Acids Res. 16:4097-4109(1988).  
 CC -1- FUNCTION: UPTAKE OF D-XYLOSE ACROSS THE BOUNDARY MEMBRANE WITH  
 CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.  
 CC -1- INDUCTION: BY XYLOSE.  
 CC -1- MISCELLANEOUS: E. COLI HAS TWO D-XYLOSE TRANSPORT SYSTEMS THAT  
 CC ACCUMULATE SUGAR AGAINST A CONCENTRATION GRADIENT: THE XYLE SYSTEM  
 CC WHICH UTILIZES THE ELECTROCHEMICAL GRADIENT OF PROTONS AND THAT IS  
 CC INSENSITIVE TO COLD OSMOTIC SHOCK AND THE XYLE SYSTEM THAT USES A  
 CC HIGH-ENERGY PHOSPHATE COMPOUND AND IS SENSITIVE TO COLD OSMOTIC  
 CC SHOCK.  
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: J02812; AAA79016.1;  
 DR EMBL: U00006; AAC43125.1;  
 DR EMBL: AE000476; AAC77001.1;  
 DR EMBL: AE000536; AAC59230.1;  
 DR EMBL: AP002568; BAB38437.1;  
 DR EMBL: X06663; CAA29863.1;  
 DR PIR: A26430; A26430.  
 DR PIR: A27418; A27418.  
 DR PIR: S00874; S00874.  
 DR EcoGene: EG1076; xyle.  
 DR InterPro: IPR003663; Sugar\_trnsportr.  
 DR InterPro: IPR003662; sub\_trnsportr.  
 DR Pfam: PF00083; sugar\_tr; 1.  
 DR PRINTS: PR00171; SUGRTNSPORT.  
 DR PROSITE: PS00216; SUGAR\_TRANSPORT\_1; 1.  
 DR PROSITE: PS00217; SUGAR\_TRANSPORT\_2; 1.  
 KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;  
 KW Complete proteome.  
 FT TRANSMEM 12 32 POTENTIAL.  
 FT TRANSMEM 57 77 POTENTIAL.  
 FT TRANSMEM 91 111 POTENTIAL.  
 FT TRANSMEM 135 155 POTENTIAL.  
 FT TRANSMEM 167 187 POTENTIAL.  
 FT TRANSMEM 202 222 POTENTIAL.  
 FT TRANSMEM 274 294 POTENTIAL.  
 FT TRANSMEM 314 334 POTENTIAL.  
 FT TRANSMEM 344 364 POTENTIAL.  
 FT TRANSMEM 371 391 POTENTIAL.  
 FT TRANSMEM 409 429 POTENTIAL.  
 FT TRANSMEM 445 465 POTENTIAL.  
 FT CONFLICT 64 64 A -> V (IN REF. 6).  
 SQ SEQUENCE 491 AA; 53608 MW; 2AF1AF9756C0B722 CRC64;

Query Match 16.6%; Score 460; DB 1; Length 491;  
 Best Local Similarity 26.1%; Pred. No. 8.9e-22;

```
Matches 144; Conservative 86; Mismatches 175; Indels 146; Gaps 14;
QY 14 VSLGGLTGYELAVISGA-----LLPLQLDFGLSCLBQEFVLGSLLLGALLASLVG 65
D 14 VSLGGLTGYELAVISGA-----LLPLQLDFGLSCLBQEFVLGSLLLGALLASLVG 65
D 16 VATLGLLGYDTAVISGTVESLNTVFAPQNLSSAANSLGFCVASALGCIIGALG 75
QY 66 GLIDCYGRKQAI-LGSNLVLLAG-----SLTGLAGSLAWLVGRVAV 107
D 66 GLIDCYGRKQAI-LGSNLVLLAG-----SLTGLAGSLAWLVGRVAV 107
D 76 GYCSNRFGRDRLKIAAVLFISGVSAMPELGFTSINPDNTVPVYLAGYVEFVYRII 135
QY 108 VGFAISLSSMACCIYVSELVGRPRGVLVSLYEAGITVGIILLSYALNYALA---GTPW-- 162
D 108 VGFAISLSSMACCIYVSELVGRPRGVLVSLYEAGITVGIILLSYALNYALA---GTPW-- 162
D 136 GGIGVGLASLSPMYIAELAPAHIRGLKLVSPNOFAIFGQLLVYCNVYFIARSGDASMLN 195
QY 163 --GWRHMFCAWAPAVLOSLSLLFLP-----AGTDETA-----THKDLIPLAGGEA 206
D 163 --GWRHMFCAWAPAVLOSLSLLFLP-----AGTDETA-----THKDLIPLAGGEA 206
D 196 TQGRYMFASECIPALLFLMLLYTVSPRWMGRGQEQAEILRKIMGNLATQAVQE 255
QY 207 PK--LGPGRPRYSFLDLFRANDNRGRTTVGLGLVLFQOLTQOPNVLCYASTIFFSVGFH 264
D 207 PK--LGPGRPRYSFLDLFRANDNRGRTTVGLGLVLFQOLTQOPNVLCYASTIFFSVGFH 264
D 256 IKHSLDGRKTKGRLMGV-----GVIVIGVMSLIFQOFVGINVLYYAPEVKTIG-A 309
QY 265 GSSAVLASVGLGAVKVAATLTANGLVDRGRARRALLAGCALMALSVSGIGLVSPVPM 324
D 265 GSSAVLASVGLGAVKVAATLTANGLVDRGRARRALLAGCALMALSVSGIGLVSPVPM 324
D 310 STDIALLOTIVGVINLTFTVLAITMDVKFKRKLQIIGALGMALGMSLGTAFY----- 364
QY 325 SGPCLAVPNATQOTGLPGDGLDSSLPPIPTNEDQREPILSTAKTKPHPRSGDPS 384
D 325 SGPCLAVPNATQOTGLPGDGLDSSLPPIPTNEDQREPILSTAKTKPHPRSGDPS 384
D 365 -----TQAPG----- 369
QY 385 APPRLASSALPGPLPARGHALLRWLTALLCLMVFSAPSFQFGPVTMLVLSIYVPEIR 444
D 385 APPRLASSALPGPLPARGHALLRWLTALLCLMVFSAPSFQFGPVTMLVLSIYVPEIR 444
D 370 -----IVALLSKLFYVAAFASWGPVYCVVLLSEIFPNAIR 404
QY 445 GRAFACNSFNWAANLIFISLSP--LD----LIGTIGLSMTFLLYGLTAVLGLGFIYLFVP 498
D 445 GRAFACNSFNWAANLIFISLSP--LD----LIGTIGLSMTFLLYGLTAVLGLGFIYLFVP 498
D 405 GRALATAVAQAOLANVFSVTFPMWDRKNSLVAHFNGFSYWIYCGMGLAALFMKKEVP 464
QY 499 ETKGSLAEID 509
D 499 ETKGSLAEID 509
D 465 ETKGKTLEBLE 475
RESULT 13
YFIG_BACSU STANDARD; PRT; 482 AA.
AC P54723;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN YFIG.
GN YFIG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96262713; PubMed=8704981;
RA Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;
RT "Detection of a 12 kb nucleotide sequence around the 76 degrees
RL region of the Bacillus subtilis chromosome.";
RL Microbiology 142:1417-1421(1996).
CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC 1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
```

```
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50343; BAA09111.1; -
DR EMBL; Z99108; CAB12655.1; -
DR Subtilist; BG11854; yfig.
DR InterPro; IPR003663; Sugar_trnsportr.
DR InterPro; IPR003662; sub_trnsportr.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PRO0171; SUGRTRNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 60 80 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 302 322 POTENTIAL.
FT TRANSMEM 332 352 POTENTIAL.
FT TRANSMEM 353 373 POTENTIAL.
FT TRANSMEM 401 421 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
SQ SEQUENCE 482 AA; 52756 MW; 90851C4F4C48EE01 CRC64;
Query Match 16.4%; Score 452.5; DB 1; Length 482;
Best Local Similarity 25.0%; Pred. No. 2.5e-21;
Matches 140; Conservative 89; Mismatches 179; Indels 151; Gaps 14;
QY 14 VSLGGLTGYELAVISGALLPLQLDFG---LSCLEQEFVLGSLLLGALLASLVGFLID 70
D 27 VSTFGGLLGYDTGVINGA-LPFMATAGQLNLTPTVEGLVASSLLGAAFGAMFGRLSD 85
QY 71 CYGRKQAIAGSLNVLGASLTGLAGSLAWLVGRVAVGFAISLSSMACCIYVSELVGP 130
D 86 RHGRKRTILYALFIAANTLCCTFSPNASVIAERFLQLAVGCASVTVPFLAEISPAE 145
QY 131 QRGVLVSLYEAGITVGIILLSYALNYALAGTPWG-----WRHMFCAWAPAVLQSL 185
D 146 RGRIVTQNELMIVIGOLLAYTFN-AIIGSTMGESANVWRYMLVIATLPVVLWFGMLIV 204
QY 186 PAG-----TDETATH--KDLIPLQGEAPKLG-----PGRPRY 216
D 205 PESPRWLAARGMGDALRVLRQIREDSQAQOEIKEIKHAEIGTAKKAGPHDFQEPWIRI 264
QY 217 SFDLFLPRARDNMRGRTTVGLGLVLFQOLTQOPNVLCYASTIFFSVGFHCGSSAVLASVGL 276
D 265 LF-----IGIAIVQOITGVNSIMYGTIELREAGFQ-TEAALIGNIAN 308
QY 277 GAVKVAATLTAMGLVDGRARRALLAG--CALMALSVSGIGLVSPVPMDSGSPCLAVPN 334
D 309 GVISVIAVFGIWLKGRVRRPMLIIGIQTWTALLIGI----- 348
QY 335 ATGOTGLPGDGLDSSLPPIPTNEDQREPILSTAKTKPHPRSGDPSAPPRALSSA 394
D 349 -----LSIV 352
QY 395 LPGPPLPARGHALLRWLTALLCLMVFSAPSFQFGPVTMLVLSIYVPEIRGRAFAFCNSP 454
D 353 LEGTP-----ALPYVWLSLITILFAFQQAISTVTLMLSEIIPPMHVRGDMGISTFC 405
QY 455 NWAANLFISSLFLDLTIGLSMTFLLYGLTAVLGLGFIYLFVPEITKGSALRIDQOPK 514
D 406 LMTANFLIGTFFPILLNHIGMSATFFIVAMNIIALFVKKYVPEITKGRSLGLEHSFRQ 465
QY 515 RRTFLSFGHR-----ONST 528
D 466 -----YGRADOEIONQT 478
RESULT 14
```

ITR1\_YEAST ITR1\_YEAST STANDARD; PRT; 584 AA.  
 AC P30605;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MYO-INOSITOL TRANSPORTER 1.  
 GN ITR1 OR YDR497C OR D9719.3.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91250431; PubMed=2040626;  
 RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;  
 RT "Isolation and characterization of two distinct myo-inositol  
 transporter genes of Saccharomyces cerevisiae.";  
 RL J. Biol. Chem. 266:11184-11191(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,  
 RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,  
 RA Hunkle-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,  
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,  
 RA Winant A., Yelton M., Botstein D., Davis R.W.;  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: MAJOR TRANSPORTER FOR MYO-INOSITOL.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; D90352; BAA14366.1; -  
 DR EMBL; U33057; AAB64939.1; -  
 DR PIR; A40538; A40538.  
 DR SGD; S0002905; ITR1.  
 DR InterPro; IPR003663; Sugar\_transport.  
 DR InterPro; IPR003662; sub\_transport.  
 DR Pfam; PF00083; sugar\_tr; 1.  
 DR PRINTS; PR00171; SUGRTRNSPT.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 2.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 DR Transmembrane; Sugar transport; Glycoprotein.  
 FT TRANSMEM 87 107 POTENTIAL.  
 FT DOMAIN 108 129 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 130 150 POTENTIAL.  
 FT TRANSMEM 159 177 POTENTIAL.  
 FT TRANSMEM 187 207 POTENTIAL.  
 FT TRANSMEM 216 236 POTENTIAL.  
 FT TRANSMEM 245 265 POTENTIAL.  
 FT DOMAIN 266 337 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 338 358 POTENTIAL.  
 FT TRANSMEM 373 393 POTENTIAL.  
 FT TRANSMEM 404 426 POTENTIAL.  
 FT TRANSMEM 446 467 POTENTIAL.  
 FT TRANSMEM 480 500 POTENTIAL.  
 FT TRANSMEM 509 529 POTENTIAL.  
 FT DOMAIN 530 584 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 371 371 N-LINKED (GLNAC. .) (POTENTIAL).  
 FT CONFLICT 43 44 TL -> HI (IN REF. 1).  
 SQ SEQUENCE 584 AA; 63569 MW; 42543E30A102DC65 CRC64;

Query Match 15.8%; Score 438; DB 1; Length 584;  
 Best Local Similarity 24.9%; Pred. No. 2.5e-20;

Matches 136; Conservative 85; Mismatches 218; Indels 108; Gaps 12;  
 QY 4 SPPVLPCLCASVSLGLLTFGYELAVISGALLPLQLDGLSCL---EQEFLVGLLLGALL 60  
 DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 DB 83 SPFIITL-TFVASISGFMFGYDGTGYISSALISITGTDLDHKLVTYGEKEIVTAATSLGALI 141  
 QY 61 ASLVGGFLIDCYGRKQAILGSLNVLVLAGSLTLGLAGSLAWLVGLRAVVGFAISLSSMACC 120  
 DB : : : : : ||||| :  
 DB 142 TSIAGTANDIFRRRCRLMGSLNMFVIGAILQVSAHTFQWQVGRIMGVGVGIGSLIAP 201  
 QY 121 IYSELVGPQRGVLSVLYEAGITVGLISLALNYALAGTPWGRHMFAGWATAPAVLOS 180  
 DB : : : : : ||||| :  
 DB 202 LFISEIAPKMRGLTVINSLWLTGGQVAYGGAGLNYNNNGWRILVGLSLPTAVQFT 261  
 QY 181 SLLELP-----AGTDETATH-----KDLPIQGGGEAPKLGPERP 214  
 DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 DB 262 CLCLFDPDTPRYVMKGLARATEVLKRSYTDTSSEIERKVEELVTLNQIPCKNVPEKV 321  
 QY 215 RYSELDLFRANDNRGRITVGLGLVLFQOLTGPQNVLCYASTIFSSVGFHGGSSAVLASV 274  
 DB : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 DB 322 WNTIKELHTVPSNLRALIIICGLQAIQOFTGWSLMYFSGTIFETVGFK-NSSAV--SI 377  
 QY 275 GLCAVKVAATLTMGLVDRAGRALLLAGC--ALMAISVSGILVGFAYPMDSGSPCLAV 332  
 DB : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 DB 378 IVSGTNEIFTLVAFESIDKIGRRITILLGLPGMTALVVCISAFHFLGKIFD-GAVAVV 436  
 QY 333 PNATGQTGLPGDGLLODSSLPPIPTWEDOREPILSTAKTKPHRSGDPSAPPRALS 392  
 DB : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 DB 437 SSGFSSWGI-----445  
 QY 393 SALPGPPLPARGHALLRWLTALLCLMVFSAFSGFGPVTWLVSEIYVEIRGAFACFN 452  
 DB : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 DB 446 -----VIIVFIIVFAAFYALGIGTVPW-QQSELPQNVRGIGTSYAT 486  
 QY 453 SFNWAANLFTSLFLDLIGTIGLSWTELLYGLTAVLGLGFIYLFVPEYKQSLAEIDQQF 512  
 DB : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 DB 487 ATNWAQSLVIASFTLWMIQNTIPAGTFAFFAGLSCLSTIFCYFCYPELSGLEEV-QTI 545  
 QY 513 QKRRFTL 519  
 DB : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 DB 546 LKDGFIN 552  
 RESULT 15  
 STA\_RICCO  
 ID STA\_RICCO STANDARD; PRT; 522 AA.  
 AC 010710;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE SUGAR CARRIER PROTEIN A.  
 GN STA.  
 OS Ricinus communis (Castor bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
 OX NCBI\_TaxID=3988;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. CARMENCITA; TISSUE=Cotyledon;  
 RA Weig A., Franz J., Sauer N., Komor E.;  
 RT "Isolation of a family of cDNA-clones from Ricinus communis L.  
 RT with close homology to the hexose carriers.";  
 RL J. Plant Physiol. 143:178-183(1994).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).  
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC or send an email to license@lsb-slb.ch).
DR EMBL; L08197; AAA79769.1;
DR InterPro; IPR003663; Sugar_transportr.
DR InterPro; IPR003662; sub_transportr.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PRO0171; SUGRTRANSPORT.
DR PRINTS; PRO0172; GLUCTRNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Transport; Sugar transport; Symport.
FT TRANSMEM 23 43 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 144 164 POTENTIAL.
FT TRANSMEM 205 225 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 322 342 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 387 407 POTENTIAL.
FT TRANSMEM 427 447 POTENTIAL.
FT TRANSMEM 452 472 POTENTIAL.
SQ SEQUENCE 522 AA; 56873 MW; 327C0F186029A586 CRC64;

Query Match 15.6% Score 430.5; DB 1; Length 522;
Best Local Similarity 25.5%; Pred. No. 6.4e-20;
Matches 145; Conservative 84; Mismatches 168; Indels 171; Gaps 18;

QY 14 VSLGLGTFYGVYELAVISGALLPLQLDFGLSCLEQEF----- 49
DB 33 VAAVGGSIFGYDIG-ISGGVISMD-----AFLEKFRSVYLLKKKHANNYCKYDDORLA 86
QY 50 -LVGSLLLGALLASVGGFLDYGRKQAILGSNLVLAGSLTGLAGSLAWLVLAGRAV 108
DB 87 AFTSSLYLAGLAASLVAGPITRIYGRASIIISGGISFLIGALNATAINLMLLLGRIML 146
QY 109 GFAISLSSMACCIYVSELVGRQGVLSYVYAGITVGLISYALYALAGT-----PWGW 164
DB 147 GVGIGFGNQAVPLVSEMAPTHLRGGLNMFQLATTSGITANMVNY--GTHKLESWG 203
QY 165 RHMFGWATAPAVLOSLSLFLP-----AGTDET-ATHKDLIPL 201
DB 204 RLSGLAALPALLMTIGLLLPETPNLSIEOGLHEKGRNVLEKIRGTHKVDADFQDML-- 261
QY 202 QGGEAPKLGPRPRYSFLDLFRARDNMRGRTVGLGLVFOOLTGPQNVLCYASTIFSSV 261
DB 262 ---DASELA-NSIKHPFNILEKRN--RPQVMAIFMPTFQILTGINILFYAPPLFQSM 315
QY 262 GFHGGSSAVLASVGLGAVKVAATLTAMGLVDRAGRALLAG-----CALMALSVSGIG 315
DB 316 GF-GGNAALYSSAVTGAVLCSSFTFISATVDRGRFLLISGGIQMITCQIVAIILGVK 374
QY 316 LVSEFVPMDSGPSCLAVPNATGTCGLPGDGLLDSSSLPIPTNEDQREPIILSTAKTK 375
DB 375 F-----GDNQQLSKS----- 384
QY 376 PPRSGDPSAPPRIALSSALPGPPLPARGHALLRWTALLCLAV--FVSAPSFSGFPVTWL 433
DB 385 -----FSLVVMICLFVLAFGWSGPIGWT 410
QY 434 VLSEIYVPEIRGRAFCNCFNWAANLF-----ISLSFLDLIGTIGLSWTFLYGLTAVLG 489
DB 411 VPSEIFPLETRSAG-----QSTIVAVNLFETTFVIAOSFPSLLCAFKFG-IFLFFAGWVTVM 465
QY 490 LGFIYLFVPETKQSLAEIDQOQKRRF 517
DB 466 TAFVYIFLPETKGVPIEEMIFLRKHWF 493
```

**THIS PAGE BLANK (USPTO)**